05-22-06

AF
ZFW

PATENT
Customer Number 22,852
Attorney Docket No. 07447.0061-00000
Xerox Ref. No.: D/A0A25

CERTIFICATE UNDER 37 CFR § 1.10 OF MAILING BY "EXPRESS MAIL"

EV 847360309 US	May 19, 2006
USPS Express Mail Label Number	Date of Deposit

I hereby certify that this correspondence is being deposited with the United States Postal Services "Express Mail Post Office to Addressee" service under 37 CFR §1.10 on the date indicated above and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

By: Mary McCauley
Mary A. McCauley

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of:	
Eric GAUSSIER et al.	
Application No.: 09/982,236	Group Art Unit: 2161
Filed: October 19, 2001	Examiner: NGUYEN, Cam Linh T.
For: METHODS, SYSTEMS AND ARTICLES OF MANUFACTURE FOR SOFT HIERARCHICAL CLUSTERING OF CO-OCCURRING OBJECTS	Confirmation No.: 7611

Mail Stop Appeal Brief--Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

TRANSMITTAL OF APPEAL BRIEF (37 C.F.R. 41.37)

Transmitted herewith is the APPEAL BRIEF in this application with respect to the Notice of Appeal filed on March 21, 2006.

This application is on behalf of

☐ Small Entity ☒ Large Entity

Pursuant to 37 C.F.R. 41.20(b)(2), the fee for filing the Appeal Brief is:

☐ $250.00 (Small Entity)

☒ $500.00 (Large Entity)

TOTAL FEE DUE:

Notice of Appeal Fee	$0.00 (previously paid on November 22, 2005)
Extension Fee (if any)	$0.00
Total Fee Due	$500.00

☒ Please charge the fee of $500.00 to our Deposit Account No. 06-0916.

PETITION FOR EXTENSION. If any extension of time is necessary for the filing of this Appeal Brief, and such extension has not otherwise been requested, such an extension is hereby requested, and the Commissioner is authorized to charge necessary fees for such an extension to our Deposit Account No. 06-0916. Please charge any additional required fees to our Deposit Account No. 06-0916.

Respectfully submitted,

FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, L.L.P.

Dated: May 19, 2006

By: _____

Venk Krishnamoorthy, Ph.D.
Reg. No. 52,490

Post Office Address (to which correspondence is to be sent)

Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
901 N.Y. Ave. N.W.
Washington, D.C. 20001-4413
(202) 408-4000

PATENT
Customer Number 22,852
Attorney Docket No. 07447.0061-00000
Xerox Ref. No.: D/A0A25

CERTIFICATE UNDER 37 CFR § 1.10 OF MAILING BY "EXPRESS MAIL"

EV 847360309 US	May 19, 2006
USPS Express Mail Label Number	Date of Deposit

I hereby certify that this correspondence is being deposited with the United States Postal Services "Express Mail Post Office to Addressee" service under 37 CFR §1.10 on the date indicated above and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

By: Mary McCauley
Mary A. McCauley

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of:	
Eric GAUSSIER et al.	
Application No.: 09/982,236	Group Art Unit: 2161
Filed: October 19, 2001	Examiner: NGUYEN, Cam Linh T.
For: METHODS, SYSTEMS AND ARTICLES OF MANUFACTURE FOR SOFT HIERARCHICAL CLUSTERING OF CO-OCCURRING OBJECTS	Confirmation No.: 7611

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450.

Sir:

APPEAL BRIEF UNDER 37 C.F.R. § 41.37

Applicants appeal the final rejections of claims 1-26 in the Office Action mailed on December 21, 2005, ("December 21 Action"). Applicants reinstated the Notice of Appeal under 37 C.F.R. § 41.31 on March 21, 2006. The Notice of Appeal and the fee set forth in 37 C.F.R. § 41.37 were originally filed on October 19, 2005, along with a request for a pre-appeal brief conference. The final rejection was withdrawn and prosecution was reopened following the decision of the pre-appeal brief conference panel. The Examiner then issued another final rejection on December 21, 2005, leading

to the present appeal. Applicants request that the Board of Appeals reverse in whole the rejections of claims 1-26 and order the allowance of these claims.

I. Real Party Interest

The real party in interest is Xerox Corporation.

II. Related Appeals and Interferences

There are no related appeals or interferences at this time.

III. Status of Claims

Claims 1-26 are pending and stand rejected. Applicants appeal the rejections of claims 1- 26.

IV. Status of Amendments

All amendments for this application have been entered.

V. Summary of Invention

The application describes methods, systems, and articles of manufacture for soft hierarchical clustering of objects based on a co-occurrence of object pairs. Clustering allows data to be hierarchically grouped (or clustered) based on its characteristics, so that objects, such as text data in documents that are similar to each other are placed in a common cluster in a hierarchy. In soft hierarchical clustering an object may be assigned to more than one cluster in a hierarchy as opposed to a hard assignment whereby an object is assigned to only one cluster in the hierarchy.

A modified Expectation-Maximization (EM) process is performed on object pairs reflecting documents and words, respectively, such that a given class of the objects ranges over all nodes of a topical hierarchy (as opposed to the leaves alone) and the assignment of a document to a topic may be based on any ancestor of the given class. Moreover, the assignment of a given document to any topic in the hierarchy may also

be based on a particular (document, word) pair under consideration during the process. The modified EM process may be performed for every child class that is generated from an ancestor class until selected constraints associated with the topical hierarchy are met. A representation of the resultant hierarchy of topical clusters may be created and made available to entities that request the topics of the document collection. See e.g. pg. 4 lines 22-23, and pg. 5, lines 1-11.

The modified algorithm eliminates the reliance on leaf nodes alone and allows any set S_i to be explained by a combination of any leaves and/or ancestor nodes included in an induced hierarchy. That is, *i* objects may not be considered as blocks, but rather as pieces that may be assigned in a hierarchy based on any *j* co-occurring objects. In one configuration, a topical clustering application performed by a computer may assign parts of a document *i* to different nodes in an induced hierarchy for different words *j* included in the document *i*. See e.g. pg. 15, lines 10-20.

For example, the probability of observing any pair of co-occurring objects, such as documents and words (*i*, *j*), may be modeled by defining a variable $I_{r\alpha}$ (controls the assignment of documents to a hierarchy) such that it is dependent on the particular document and word pair (*i*, *j*) under consideration during a topical clustering process. In one configuration, the class α may range over all nodes in an induced hierarchy in order to assign a document (*i* object) to any node in the hierarchy, not just leaves. Furthermore, by defining a class *v* as any ancestor of α in the hierarchy the nodes may be hierarchically organized. See e.g. pg. 15, lines 21-23, and pg. 16, lines 1-6.

Different *j* objects may be generated from different vertical paths of an induced hierarchy. That is, from paths in the hierarchy associated with non null values of $I_{i\alpha}$.

Furthermore, because α may be any node in the hierarchy, the *i* objects may be assigned to different levels of the hierarchy. Accordingly, implementation of the model results in a pure soft hierarchical clustering of both *i* and *j* objects by eliminating any hard assignments of these objects. See, e.g., pg. 18, lines 10-21.

The model may be implemented for a variety of applications, depending upon the meaning given to objects *i* and *j*. For example, it may be applied to document clustering based on topic detection. In such a configuration, *i* objects may represent documents and *j* objects may represent words included in the documents. Clusters or topics of documents may be represented by leaves and/or nodes of an induced hierarchy. The topics associated with the document collection may be obtained by interpreting any cluster as a topic defined by the word probability distributions, $p(j \mid v)$. The soft hierarchical model may take into account several properties when interpreting the clusters, such as: (1) a document may cover (or be explained by) several topics (soft assignment of *i* objects provided by the probability $p(i \mid \alpha)$); (2) a topic is best described by a set of words, which may belong to different topics due to polysemy (the property of a word to exhibit several different, but related meanings) and specialization (soft assignment of *j* objects provided by the probability $p(j \mid v)$); and (3) topics may be hierarchically organized, which corresponds to the hierarchy induced over clusters. See, e.g., pg. 20, lines 25-30, and pg. 21, lines 1-11.

One or more conditions associated with a hierarchy that may be induced may allow a computer to determine when an induced hierarchy reaches a desired structure with respect to the clusters defined therein. For example, a condition may be defined that instructs a processor to stop locating co-occurring objects (*i*, *j*) in a document

collection that is being clustered based on a predetermined number of leaves, and/or a level of the induced hierarchy. See, e.g., pg. 23, lines 1-11.

Pending independent claim 1 recites a method performed by a computer for clustering a plurality of documents in a structure comprised of a plurality of clusters hierarchically organized, wherein each document includes a plurality of words and is represented as a set of (document, word) pairs, the method comprising: accessing the document collection; performing a clustering process that creates a hierarchy of clusters that reflects a segregation of the documents in the collection based on the words included in the documents, wherein any document in the collection may be assigned to a first cluster in the hierarchy based on a first segment of the respective document, and the respective document may be assigned to a second cluster in the hierarchy based on a second segment of the respective document, wherein the first and second clusters are associated with different paths of the hierarchy; storing a representation of the hierarchy of clusters in a memory; and making the representation available to an entity in response to a request associated with the document collection. Claims 2-7 all ultimately depend from claim 1.

Pending independent claim 8 recites a method performed by a computer for determining topics of a document collection, the method comprising accessing the document collection, each document including a plurality of words and being represented as a set of (document, word) pairs; performing a clustering process including: creating a tree of nodes that represent topics associated with the document collection based on the words in the document collection, wherein any node in the tree may include a word that is shared by another node in the tree, and assigning fragments

of one or more documents included in the document collection to multiple nodes in the tree based on the (document, word) pairs; storing a representation of the tree in a memory; and making the representation available for processing operations associated with the document collection. Claim 9 ultimately depends from claim 8.

Pending independent claim 10 recites a method performed by a processor for clustering data in a database, the method comprising: receiving a collection of documents, each document including a plurality of words and being represented as a set of (document, word) pairs; creating a first ancestor node reflecting a first topic based on words included in the collection of documents; creating descendant nodes from the first ancestor node, each descendant node reflecting descendant topics based on the first node, until a set of leaf nodes reflecting leaf topics are created. The step of creating descendant nodes includes assigning each document in the collection to a plurality of descendant and leaf nodes; and providing a set of topics associated with the collection of documents based on the created nodes and assignment of documents, wherein the descendant and leaf nodes may be created based on one or more words included in more than one document in the collection of documents. Claim 11 ultimately depends from claim 10.

Pending independent claim 12 recites a method performed by a processor for clustering data in a database, the method comprising: receiving a collection of documents, each document including a plurality of words and being represented as a set of (document, word) pairs; creating a hierarchy of nodes based on the words in the collection of documents, each node reflecting a topic associated with the documents, wherein the hierarchy of nodes includes ancestor nodes, descendant nodes, and leaf

nodes; assigning each document in the collection to a plurality of nodes in the hierarchy, wherein each document may be assigned to any of the ancestor, descendant, and leaf nodes; and providing a set of topic clusters associated with the collection of documents based on the created nodes and assignment of documents, wherein the hierarchy may include a plurality of nodes that are each created based on a same set of words included in the collection of documents.

Pending independent claim 13 recites a method performed by a computer for clustering data stored on a computer-readable medium, the method comprising: receiving a collection of data objects, represented as a set of (first data object, second data object) pairs; for each first data object: assigning the first data object to a first node in a hierarchy of nodes based on the second data objects included in the first data object, wherein the first node may be any node included in the hierarchy and wherein two or more nodes in the hierarchy may share the same second object; creating a final hierarchy of nodes arranged in clusters based on the assignment of the first data objects; storing a representation of the final hierarchy in a memory; and making the representation of the final hierarchy available to an entity in response to a request associated with the collection of first data objects.

Pending independent claim 14 recites a method performed by a processor for clustering data in a database, the method comprising: receiving a request from a requesting entity to determine topics associated with a collection of documents, each document including a plurality of words and being represented as a set of (document, word) pairs; determining the topics associated with the collection of documents based on a hierarchy including a plurality of clusters, wherein each cluster reflects a topic and

a document in the collection may be assigned to a set of clusters in the hierarchy based on different words included in the document, and wherein each cluster in the set may be associated with different paths in the hierarchy; storing a representation of the hierarchy in a memory; and making the representation available to the requesting entity.

Pending independent claim 15 recites a computer-implemented method for clustering a plurality of multi-word documents into a hierarchical data structure including a root node associated with a plurality of sub-nodes, wherein each sub-node is associated with a topic cluster based on the plurality of documents, the method comprising: retrieving a first document; associating the first document with a first topic cluster based on a first portion of the first document; associating the first document with a second topic cluster based on a second portion of the document; and providing a representation of topics associated with the plurality of multi-word documents based on the hierarchical data structure including the first and second topic clusters, wherein the first and second topic clusters are associated with a different sub-node. Claims 16-19 all ultimately depend from claim 15.

Pending independent claim 20 recites a computer-implemented method for clustering data reflecting users, represented as a set of (data, user) pairs, into a hierarchical data structure including a root node associated with a plurality of sub-nodes, wherein each sub-node represents an action that is performed on a document collection, comprising: accessing a user data collection reflecting a plurality of users who each perform at least one action on the document collection, wherein each action may be unique; performing a clustering process that creates the hierarchical data structure, wherein the clustering processing comprises: retrieving a first user data,

associated with a first user, from the user data collection, associating the first user data with a first sub-node based on a first action performed by the first user on the document collection, and associating the first user data with a second sub-node provided the first user data is based on a second action, wherein the first and second sub-nodes are associated with different descendent paths of the hierarchical data structure; storing a representation of the hierarchical data structure in a memory; and making the representation available to an entity in response to a request associated with the user data collection. Claim 21 ultimately depends from claim 20.

Pending independent claim 22 recites a computer-implemented method for clustering a plurality of images based on text associated with the images, where each image is represented as a set of pairs (image, image feature) and (image, text feature), into a hierarchical data structure including a root node associated with a plurality of sub-nodes, wherein each sub-node represents a different topic, the method comprising: accessing an image collection; performing a clustering process that creates the hierarchical data structure, wherein the clustering processing comprises: associating a first image with a first sub-node based on a first portion of text associated with the first image, and associating the first image with a second sub-node based on a second portion of text associated with the first image, wherein the first and second sub-nodes are associated with different descendant paths of the hierarchical data structure; storing a representation of the hierarchical data structure in a memory; and making the representation available to an entity in response to a request associated with the image collection.

Pending independent claim 23 recites a computer-implemented method for clustering customer purchases, represented as a set of (customer, purchase) pairs, into a hierarchical data structure including a root node associated with a plurality of sub-nodes, wherein each sub-node represents a group of customers who purchased the same type of product from one or more business entities, the method comprising: accessing information associated with a plurality of customers who purchased various types of products from a plurality of business entities; performing a clustering process that creates the hierarchical data structure, wherein the clustering processing comprises: associating a first customer with a first sub-node based on a first type of product purchased from a first business entity, and associating the first customer with a second sub-node provided the first customer is based on a second type of product that the first customer purchased from a second business entity, wherein the first and second sub-nodes are associated with different descendant paths of the hierarchical data structure; storing a representation of the hierarchical data structure in a memory; and making the representation available in response to a request associated with the customer data collection. Claims 24-26 all ultimately depend from claim 23.

VI. Grounds of Rejection to be Reviewed on Appeal

A. Whether the claims 1, 8, 10, 12-16, 13, 14, 15, 20-23 should be rejected under 35 U.S.C. § 103(a) as unpatentable in light of U.S. Patent No. 5,761,418 to Francis et. al. ("Francis") in view of U.S. 6,078,943 to Aoki et. al. ("Aoki). .

B. Whether claims 2-7, 9, 11, 17-19, and 24-26 should be rejected under 35 U.S.C. § 103(a) as unpatentable in light of Francis in view of Aoki and further in view of U.S. Patent No. 6,233,575 to Agrawal et. al. ("Agrawal").

VII. Argument

In the Final Office Action mailed Dec. 21, 2005, ("Final Office Action") the Examiner rejected independent claims 1, 8, 10, 12, 13, 14, 15, 20, 22, and 23 under 35 U.S.C. § 103(a) as unpatentable in light of Francis in view of Aoki. Claims 16 and 21, which depend from claims 15 and 20, respectively were also rejected under 35 U.S.C. § 103(a) as unpatentable in light of Francis in view of Aoki.

In addition, the Examiner rejected claim 11, which depends from independent claim 10, under 35 U.S.C. § 103(a) as being unpatentable in light of Francis in view of Aoki and further in view of Agrawal. Claims 2-7, 9, 17-19, and 24-26, which depend from independent claims 1, 8, 15, and 23, respectively, were also rejected under 35 U.S.C. 103(a) as being unpatentable in light of Francis in view of Aoki and further in view of Agrawal et al (U.S. 6,233,575).

As Appellants will show below, Francis does not disclose the creation or use of any hierarchical data structures to store representations of hierarchies of resource clusters. Indeed, the organization recited in Francis teaches away from the use of such structures to promote good scalability as the number of resources, terms, and term combinations grow. The teachings in Francis are directed to methods to obviate the need for such structures so that no resource is associated with any explicit information about what cluster it belongs to, nor is there a need for such information to exist anywhere. In fact, Francis mentions cluster hierarchies only to distinguish them from the methods it adopts. Because Francis specifically states that the methods it adopts are significantly different from hierarchical cluster-based systems, one skilled in the art would lack the motivation to modify Francis in an attempt to achieve the claimed

invention. Fatal to the Examiner's analysis, therefore, is the fact that Francis teaches away from and uses a radically different approach from methods outlined in the invention and in Aoki. Moreover, a combination of Francis and Aoki would not be practicable, and their teachings, if combined, would not yield the present invention.

In summary, since the Examiner's arguments fail to establish a *prima facie* case of obviousness, Appellants respectfully submit that the rejection of all claims under 35 U.S.C. § 103 is improper.

A. The Test to Establish a Prima Facie Case of Obviousness

The case law and the M.P.E.P. clearly set forth the requirements to establish a *prima facie* case of obviousness, and both place the burden of doing so squarely on the Examiner. To establish a *prima facie* case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations. The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art, not in applicant's disclosure. *See* M.P.E.P. §2143 (citing *In re* Vaeck, 947 F.2d 488, 20 U.S.P.Q.2d 1438 (Fed. Cir. 1991)).

B. Claims 1, 8, 10, 12, 13, 14, 15-16, and 20-23 Are Patentable over Francis in View of Aoki

In establishing a *prima facie* case of obviousness, the Examiner must first demonstrate teaching or motivation in the references themselves to suggest the

combination. Indeed, no such teachings or motivation have been shown. Further, because Francis teaches away from the use of methods outlined in the invention and in Aoki, a practitioner would have no reasonable expectation of success in effecting a combination of Francis and Aoki to achieve the invention. Moreover, such a combination would not be practicable. Finally, the Examiner fails to show how the cited references teach or suggest all claim limitations. In attempting to satisfy this requirement, the Examiner strings together passages scattered throughout Francis and provides interpretations of drawings in Francis that purportedly teach some of the elements of claim 1. The Examiner then attempts to cure admitted deficiencies in Francis by combining it with Aoki, but fails to specifically point out how Aoki cures the deficiencies.

The Examiner has also failed to specifically address how the teachings in Francis and Aoki, either individually, or in combination teach the limitations of independent claims 8, 10, 12, 13, 14, 15, 20, 22, and 23. The Examiner has also not discussed or pointed out how elements in the above claims correspond or are analogous to elements in claim 1, which is the only independent claim specifically addressed in the Final Office Action.

1. Examiner Fails to Show Motivation to Combine

The Examiner fails to show any teaching or motivation in the references themselves to suggest the combination. Indeed, Examiner's sole basis for combining the references is because Francis discusses prior art where the resources may be organized in a hierarchical order. See Final Office Action, pg. 6, lines 21-22; pg. 7, lines 1-5. However, Francis mentions cluster hierarchies only to distinguish them from the

methods it adopts. Because Francis specifically states that the methods it adopts are significantly different from hierarchical cluster-based systems, one skilled in the art would lack the motivation to modify Francis to in an attempt to achieve the claimed invention. See, Francis, col. 7, lines 13-25.

Teaching away pertains to proposed modifications that render the prior art unsatisfactory for its intended purpose or that change the principle of operation of a reference. A prior art reference that teaches away from the claimed invention is a significant factor to be considered in determining obviousness. *See*, M.P.E.P. § 2145 (X)(D) (8th Ed., Rev. 3, Aug. 2005). By failing to correctly consider and substantively address Appellant's arguments that Francis teaches away from the methods described in the invention and in Aoki, Examiner has made a flawed rejection of the claims.

Francis is directed to distributed topology creation and maintenance, where a resource does not need to have associated with it any explicit information about what cluster it belongs to, nor does it require such information to exist anywhere. According to Francis, this lack of explicit information also contributes to good scaling, especially in the case where there are a large number of resources, each of which contains a large number of terms or term combinations. See, Francis, col. 6, lines 54-66; col. 7 lines 13-25; col. 9, lines 11-44. Accordingly, the teachings in Francis are directed to methods to obviate the need for such structures so that no resource is associated with any explicit information about what cluster it belongs to, nor is there a need for such information to exist anywhere., i.e. the clusters are latent or notional. By obviating the need to store such information centrally, Francis limits the latency of searches as the number of resources grows. In contrast, Aoki requires a centralized cluster database that stores a

cluster of node information elements in a tree structure based degree of similarity in all of the documents. Therefore, Francis explicitly teaches away from the use of centralized indexing or centralized database methods for searching and/or navigation such as the methods outlined in Aoki, in which "a cluster database storing a cluster of node information elements" in a "tree structure based degree of similarity in all of the documents" is used. See Aoki, Abstract.

Examiner has failed to address Appellant's arguments that Francis teaches away from methods outlined in Aoki. The Examiner's premise that references may be combined solely on the basis that there is a reference to some methods in Francis, without a corresponding analysis of those teachings in Francis is without foundation and lacking in statutory or case-law support. Therefore, the Examiner's basis for the obviousness rejection is clearly erroneous because the purported basis for rejection arises as a consequence of the defective combination of Francis and Aoki.

2. No Reasonable Expectation of Success Because Francis Teaches Away From the Combination

By failing to substantively address Appellants' arguments that Francis teaches away from the methods described in Aoki, the Examiner has also failed to show that a person of reasonable skill in the art effecting such a combination would have a reasonable expectation of success.

As noted above, Francis teaches away from the methods outlined in Aoki. In Francis resource information is distributed and nodes have sparse links to similar resources and any cluster information is latent to promote scalability. By contrast, in Aoki an approximation of a degree of similarity is used to make hard assignments of

documents to a cluster that is stored centrally in a cluster database. Therefore, a combination of the teachings in Francis and Aoki as suggested by the Examiner would not be practicable and a person of reasonable skill in the art would have no reasonable expectation of success in effecting such a combination in an attempt to achieve the present invention. Moreover, such a combination of the references would neither anticipate nor yield the present invention.

3. Examiner Fails to Establish That Francis and Aoki in Combination Teach All Claim Limitations

The Examiner starts by broadly stating that Francis "discloses a method for clustering a plurality of documents comprised of a plurality of clusters wherein the clustering process creates a hierarchy of clusters. See, Final Office Action, pg. 3, lines 1-13, Final Office Action. In the Examiner's opinion the scattered passages and drawings (citing Francis, Fig. 1 and col. 7, lines 4-31; Fig. 5 and col. 13, lines 45-51;) show a hierarchical structure of clusters.

Even a cursory reading of the selected passages would show that Examiner's interpretations do not find support in Francis. As recited in Francis, Fig. 1 "shows a number of resources organized into a sparsely-connected *mesh* network" (See, Francis, col. 7, lines 5-6. See also col. 6, lines 66-67). Contrary to the Examiner's assertion, a sparsely connected mesh network is not hierarchically organized. A conventional mesh structure does not provide or suggest a hierarchy and no hierarchical structure is suggested in Francis. Indeed, the only mentions of hierarchy or hierarchical organization in Francis serve to distinguish prior methods from those presented in Francis.

The Examiner concedes in a convoluted statement that Francis "does not clearly teach that the clusters are associated with different paths of the hierarchy or the plurality of clusters hierarchical organized, wherein each document includes a plurality of words and is represented as a set of (document, word) pairs." See, Final Office Action, pg. 4, lines 3-5. The Examiner then summarily leaps to combine the teachings in Francis with Aoki, and relies on the Abstract of Aoki to make rejections. However, the Examiner fails to show how the addition of Aoki cures admitted deficiencies in Francis. Specifically, the Examiner fails to show at least how Francis and Aoki individually, or in combination, teach that clusters may be associated with different paths of the hierarchy, as recited in claim 1.

Indeed, Aoki teaches the "hard" assignment of documents to clusters organized in a tree structure so that a document can only belong to one cluster. The method disclosed in Aoki, "selects a specific number of documents, clusters those, assigns the remaining *non-selected documents* respectively to a leaf node to be similar to the documents in the cluster, and repeats recursively the above operations toward a direction of the leaf node of cluster." See, Aoki, Abstract. The algorithm disclosed in Aoki clusters selected documents, and assigns any non-selected documents to a current leaf node. The algorithm is then recursively applied to the current leaf node to generate new leaf nodes. Therefore, a document is associated with only one cluster and a cluster can only be associated with one path in the hierarchy, because the recursive clustering is only applied to *non-selected* documents associated with a current (unclustered) leaf node. See, e.g., Aoki, col. 3, lines 43-49; col. 4, lines 5-15; col. 8 lines 7-32.

Despite a recognition of a duty to address each and every element of the pending claims, the Examiner has failed to address how the teachings in Francis and Aoki, either individually, or in combination teach the limitations of independent claims 8, 10, 12, 13, 14, 15, 20, 22, and 23. Further, the Examiner has not pointed out how elements in the above claims correspond or are analogous to elements in claim 1. In an apparent attempt to explain the failure to address the limitations of the above claims, the Examiner states "it is logical for the examiner to focus on the limitations that are 'crux of the invention' and not involve a lot of energy and time for the things that are not central to the invention, but peripheral." See, Final Office Action, pg. 7, lines 7-9. Appellants disagree with Examiner's comments regarding the duty to address each and every element of the pending claims and any characterizations of the "crux of the invention." However, notwithstanding Examiner's failure to address limitations in each of the above claims individually, Appellants provide at least the following reasons for their patentability over the cited art. Appellants reserve the right to further refine and advance arguments based on Examiner's response.

Francis and Aoki, either individually, or in combination, do not teach or suggest at least the process of

> performing a clustering process including: creating a tree of nodes that represent topics associated with the document collection based on the words in the document collection, wherein any node in the tree may include a word that is shared by another node in the tree, and assigning fragments of one or more documents included in the document collection to multiple nodes in the tree based on the (document, word) pairs;

as recited in claim 8.

Francis and Aoki, either individually, or in combination, do not teach or suggest at least the process of creating descendant nodes, which includes

assigning each document in the collection to a plurality of descendant and leaf nodes; and providing a set of topics associated with the collection of documents based on the created nodes and assignment of documents, wherein the descendant and leaf nodes may be created based on one or more words included in more than one document in the collection of documents

as recited in claim 10.

Francis and Aoki, either individually, or in combination, do not teach or suggest at least the process of

assigning each document in the collection to a plurality of nodes in the hierarchy, wherein each document may be assigned to any of the ancestor, descendant, and leaf nodes; and providing a set of topic clusters associated with the collection of documents based on the created nodes and assignment of documents, wherein the hierarchy may include a plurality of nodes that are each created based on a same set of words included in the collection of documents

as recited in claim 12.

Francis and Aoki, either individually, or in combination, do not teach or suggest at least the process of

assigning the first data object to a first node in a hierarchy of nodes based on the second data objects included in the first data object, wherein the first node may be any node included in the hierarchy and wherein two or more nodes in the hierarchy may share the same second object; creating a final hierarchy of nodes arranged in clusters based on the assignment of the first data objects

as recited in claim 13.

Francis and Aoki, either individually, or in combination, do not teach or suggest at least the process of

determining the topics associated with the collection of documents based on a hierarchy including a plurality of clusters, wherein each cluster reflects a topic and a document in the collection may be assigned to a set of clusters in the hierarchy based on different words included in the document, and wherein each cluster in the set may be associated with different paths in the hierarchy

as recited in claim 14.

Francis and Aoki, either individually, or in combination, do not teach or suggest at least the process of

> providing a representation of topics associated with the plurality of multi-word documents based on the hierarchical data structure including the first and second topic clusters, wherein the first and second topic clusters are associated with a different sub-node

as recited in claim 15.

Francis and Aoki, either individually, or in combination, do not teach or suggest at least the process of

> associating the first user data with a second sub-node provided the first user data is based on a second action, wherein the first and second sub-nodes are associated with different descendent paths of the hierarchical data structure

as recited in claim 20.

Francis and Aoki, either individually, or in combination, do not teach or suggest at least the process of

> associating the first image with a second sub-node based on a second portion of text associated with the first image, wherein the first and second sub-nodes are associated with different descendant paths of the hierarchical data structure

as recited in claim 22.

Francis and Aoki, either individually, or in combination, do not teach or suggest at least the process of

> associating the first customer with a second sub-node provided the first customer is based on a second type of product that the first customer purchased from a second business entity, wherein the first and second sub-nodes are associated with different descendant paths of the hierarchical data structure

as recited in claim 23.

Therefore claims 1, 8, 10, 12, 13, 14, 15, 20, 22, and 23 are patentable over Francis and Aoki either individually, or in combination.

Claim 16 depends from claim 15 and is patentable for at least the same reasons as is claim 15.

Claim 21 depends from claim 20 and is patentable for at least the same reasons as is claim 20.

C. Claims 2-7, 9, 11, 17-19, and 24-26 Are Patentable over Francis in View of Aoki further in View of Agrawal

Without repeating the arguments set forth above for claims 1, 8, 10, 12-16, and 20-23, Appellants contend that claims 2-7, 9, 11, 17-19, and 24-26 are patentable over Francis in view of Aoki further in view of Agrawal. First, the Examiner has failed to sufficiently establish motivation to combine Francis with Aoki to anticipate the claims. Second, the Examiner fails to establish likelihood of success, since Francis teaches away from methods outlined in Aoki and the claimed invention. Finally, prima facie obviousness is not established because the references individually, or in combination, do not recite all of the elements of the claims 2-7, 9, 11, 17-19, and 24-26.

In view of the clear errors in the Examiner's rejections, and omissions of one or more essential elements needed to support a prima facie case of rejection under 35 U.S.C. § 103(a), Appellants submit that the rejections of independent claims 1, 8, 10, 12-15, and 20, 22, and 23 were improper. Accordingly, Appellants request withdrawal of the rejections and allowance of these claims. Appellants submit further that dependent claims 2-7, 9, 11, 17-19, and 24-26 are also allowable because they depend

from independent claims 1, 8, and 15, respectively, which, as Appellant has pointed out previously, are allowable.

IX. Conclusion

A fundamental rule of patent examining procedure is that the burden is on the Examiner to establish at least a *prima facie* showing of obviousness before any claim can be properly rejected under 35 U.S.C. § 103. The Examiner has failed to make such a showing in this case.

The Examiner has failed to show a motivation to combine the Francis and Aoki references. Moreover, by failing to substantively address Appellant's arguments that Francis teaches away from the methods described in the invention and in Aoki, the Examiner has also failed to show that a person of reasonable skill in the art effecting such a combination would have a reasonable expectation of success. Finally, the Examiner has also failed to provide any support to show that the combination of Francis and Aoki would in fact yield the present invention.

Thus, Examiner has clearly failed to meet the burden of establishing a prima facie case of obviousness. For the foregoing reasons, Appellants respectfully request reversal of all of the bases for rejection set forth in the Grounds of Rejection to be Reviewed on Appeal section above (i.e., Section VI, items (A) – (B)) and allowance of all pending claims.

To the extent any further extension of time under 37 C.F.R. § 1.136 is required to obtain entry of this Appeal Brief, such extension is hereby respectfully requested. Please grant any extensions of time required to enter this paper and charge any additional required fees to our Deposit Account No. 06-0916.

Respectfully submitted,

FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, L.L.P.

Dated: May 19, 2006

By: 

Venk Krishnamoorthy, Ph.D.
Reg. No. 52,490

Post Office Address (to which correspondence is to be sent)

Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
901 N.Y. Ave. N.W.
Washington, D.C. 20001-4413
(202) 408-4000

Claims Appendix

Pending claims on appeal:

1. A method performed by a computer for clustering a plurality of documents in a structure comprised of a plurality of clusters hierarchically organized, wherein each document includes a plurality of words and is represented as a set of (document, word) pairs, the method comprising:

accessing the document collection;

performing a clustering process that creates a hierarchy of clusters that reflects a segregation of the documents in the collection based on the words included in the documents, wherein any document in the collection may be assigned to a first cluster in the hierarchy based on a first segment of the respective document, and the respective document may be assigned to a second cluster in the hierarchy based on a second segment of the respective document, wherein the first and second clusters are associated with different paths of the hierarchy;

storing a representation of the hierarchy of clusters in a memory; and

making the representation available to an entity in response to a request associated with the document collection.

2. The method of claim 1, wherein performing a clustering process comprises:

assigning the document collection to a first class;

setting a probability parameter to an initial value; and

determining, for each document in the collection at the value of the parameter, a probability of an assignment of the document in the collection to a cluster in the hierarchy based on a word included in the document and the first class.

3. The method of claim 2, wherein the step of determining further comprises:

determining whether the first class has split into two child classes, wherein each child class reflects a cluster descendant from an initial cluster reflected by the first class; and

increasing the value of the parameter based on the determination whether the first class has split into two child classes.

4. The method of claim 3, further comprising:

repeating the step of determining, for each document in the collection at the value of the parameter, and the step of increasing the value of the parameter until the first class has split into two child classes.

5. The method of claim 4, further comprising:

performing the clustering process for each child class until each of the respective child class splits into two new child classes reflecting clusters descendant from the respective child class.

6. The method of claim 5, further comprising:

repeating the clustering process for each new child class such that a hierarchy of clusters is created, until a predetermined condition associated with the hierarchy is met.

7. The method of claim 6, wherein the predetermined condition is one of a maximum number of leaves associated with the hierarchy and depth level of the hierarchy.

8. A method performed by a computer for determining topics of a document collection, the method comprising:

accessing the document collection, each document including a plurality of words and being represented as a set of (document, word) pairs;

performing a clustering process including:

creating a tree of nodes that represent topics associated with the document collection based on the words in the document collection, wherein any node in the tree may include a word that is shared by another node in the tree, and

assigning fragments of one or more documents included in the document collection to multiple nodes in the tree based on the (document, word) pairs;

storing a representation of the tree in a memory; and

making the representation available for processing operations associated with the document collection.

9. The method of claim 8, wherein the step of assigning comprises:

associating a set of documents in the document collection with a first class reflecting all of the nodes in the tree, wherein the set of documents may include all or some of the documents in the collection;

defining a second class reflecting any ancestor node of a node in the first class;

determining, for each document in the set, a probability that different words included in a respective document co-occurs with the respective document in any node in the tree based on the first and second classes; and

assigning one or more fragments of any document in the set to any node in the tree based on the probability.

10. A method performed by a processor for clustering data in a database, the method comprising:

receiving a collection of documents, each document including a plurality of words and being represented as a set of (document, word) pairs;

creating a first ancestor node reflecting a first topic based on words included in the collection of documents;

creating descendant nodes from the first ancestor node, each descendant node reflecting descendant topics based on the first node, until a set of leaf nodes reflecting leaf topics are created,

wherein creating descendant nodes includes:

assigning each document in the collection to a plurality of descendant and leaf nodes; and

providing a set of topics associated with the collection of documents based on the created nodes and assignment of documents,

wherein the descendant and leaf nodes may be created based on one or more words included in more than one document in the collection of documents.

11. The method of claim 10, wherein the step of creating descendant nodes comprises:

selecting a first document in the collection;

defining a first class that includes all of the nodes;

defining a second class that may include any ancestor node of any node included in the first class; and

determining, for each document in the collection, a target word of an object pair including a target document and the target word such that the first document equals the target document in the object pair based on a probability associated with the first and second classes; and

assigning the first document to any ancestor, descendant, and leaf node based on the determining.

12. A method performed by a processor for clustering data in a database, the method comprising:

receiving a collection of documents, each document including a plurality of words and being represented as a set of (document, word) pairs;

creating a hierarchy of nodes based on the words in the collection of documents, each node reflecting a topic associated with the documents, wherein the hierarchy of nodes includes ancestor nodes, descendant nodes, and leaf nodes;

assigning each document in the collection to a plurality of nodes in the hierarchy, wherein each document may be assigned to any of the ancestor, descendant, and leaf nodes; and

providing a set of topic clusters associated with the collection of documents based on the created nodes and assignment of documents,

wherein the hierarchy may include a plurality of nodes that are each created based on a same set of words included in the collection of documents.

13. A method performed by a computer for clustering data stored on a computer-readable medium, the method comprising:

receiving a collection of data objects, represented as a set of (first data object, second data object) pairs;

for each first data object:

assigning the first data object to a first node in a hierarchy of nodes based on the second data objects included in the first data object, wherein the first node may be any node included in the hierarchy and wherein two or more nodes in the hierarchy may share the same second object;

creating a final hierarchy of nodes arranged in clusters based on the assignment of the first data objects;

storing a representation of the final hierarchy in a memory; and

making the representation of the final hierarchy available to an entity in response to a request associated with the collection of first data objects.

14. A method performed by a processor for clustering data in a database, the method comprising:

receiving a request from a requesting entity to determine topics associated with a collection of documents, each document including a plurality of words and being represented as a set of (document, word) pairs;

determining the topics associated with the collection of documents based on a hierarchy including a plurality of clusters, wherein each cluster reflects a topic and a document in the collection may be assigned to a set of clusters in the hierarchy based on different words included in the document, and wherein each cluster in the set may be associated with different paths in the hierarchy;

storing a representation of the hierarchy in a memory; and

making the representation available to the requesting entity.

15. A computer-implemented method for clustering a plurality of multi-word documents into a hierarchical data structure including a root node associated with a plurality of sub-nodes, wherein each sub-node is associated with a topic cluster based on the plurality of documents, the method comprising:

retrieving a first document;

associating the first document with a first topic cluster based on a first portion of the first document;

associating the first document with a second topic cluster based on a second portion of the document; and

providing a representation of topics associated with the plurality of multi-word documents based on the hierarchical data structure including the first and second topic clusters,

wherein the first and second topic clusters are associated with a different sub-node.

16. The method of claim 15, wherein the first and second portions contain at least one unique word.

17. The method of claim 15, wherein associating the first document with a first topic cluster comprises:

assigning the plurality of multi-word documents to a first class;

setting a probability parameter to an initial value; and

determining, for the first document at the value of the parameter, a probability of an assignment of the first document to the first topic cluster based on a word included in the first document and the first class.

18. The method of claim 15, wherein associating the first document with a second topic cluster comprises:

assigning the plurality of multi-word documents to a first class;

setting a probability parameter to an initial value; and

determining a probability of an assignment of the first document to the second topic cluster based on a word included in the first document and the first class.

19. The method of claim 15, wherein providing a representation comprises:

providing the representation after each document in the plurality of multi-word documents has been associated with to at least one topic cluster corresponding to a sub-node in the hierarchy, wherein any of the plurality of multi-word documents may be

associated to more than one topic cluster based on different portions of the respective document.

20. A computer-implemented method for clustering data reflecting users, represented as a set of (data, user) pairs, into a hierarchical data structure including a root node associated with a plurality of sub-nodes, wherein each sub-node represents an action that is performed on a document collection, comprising:

accessing a user data collection reflecting a plurality of users who each perform at least one action on the document collection, wherein each action may be unique;

performing a clustering process that creates the hierarchical data structure, wherein the clustering processing comprises:

retrieving a first user data, associated with a first user, from the user data collection,

associating the first user data with a first sub-node based on a first action performed by the first user on the document collection, and

associating the first user data with a second sub-node provided the first user data is based on a second action, wherein the first and second sub-nodes are associated with different descendent paths of the hierarchical data structure;

storing a representation of the hierarchical data structure in a memory; and

making the representation available to an entity in response to a request associated with the user data collection.

21. The method of claim 20, wherein each action in the one or more actions includes:

writing to, printing, and browsing the document collection.

22. A computer-implemented method for clustering a plurality of images based on text associated with the images, where each image is represented as a set of pairs (image, image feature) and (image, text feature), into a hierarchical data structure including a root node associated with a plurality of sub-nodes, wherein each sub-node represents a different topic, the method comprising:

accessing an image collection;

performing a clustering process that creates the hierarchical data structure, wherein the clustering processing comprises:

associating a first image with a first sub-node based on a first portion of text associated with the first image, and

associating the first image with a second sub-node based on a second portion of text associated with the first image, wherein the first and second sub-nodes are associated with different descendant paths of the hierarchical data structure;

storing a representation of the hierarchical data structure in a memory; and

making the representation available to an entity in response to a request associated with the image collection.

23. A computer-implemented method for clustering customer purchases, represented as a set of (customer, purchase) pairs, into a hierarchical data structure including a root node associated with a plurality of sub-nodes, wherein each sub-node represents a group of customers who purchased the same type of product from one or more business entities, the method comprising:

accessing information associated with a plurality of customers who purchased various types of products from a plurality of business entities;

performing a clustering process that creates the hierarchical data structure, wherein the clustering processing comprises:

associating a first customer with a first sub-node based on a first type of product purchased from a first business entity, and

associating the first customer with a second sub-node provided the first customer is based on a second type of product that the first customer purchased from a second business entity, wherein the first and second sub-nodes are associated with different descendant paths of the hierarchical data structure;

storing a representation of the hierarchical data structure in a memory; and

making the representation available in response to a request associated with the customer data collection.

24. The method of claim 1, wherein the representation defines the probability of a document as the product of the probability of the (document, word) pairs it contains.

25. The method claim 24, wherein the product is calculated after mixing the document-word pairs over the clusters.

26. The method claim 25, wherein mixing the (document, word) pairs over the clusters comprises a probability model of the form:

$$P(x) = \sum_{c} P(c)P(x \mid c)$$

wherein *c* is the group of clusters involved in the calculation, and *x* is a (document, word) pair.

Evidence Appendix

None.

Related Proceedings Appendix

None.